

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

АДАМОВ ОЛЕКСАНДР СЕМЕНОВИЧ

УДК 658:512.011: 681.326: 519.713

**МОДЕЛІ І МЕТОДИ ЗАХИСТУ КІБЕРПРОСТОРУ
НА ОСНОВІ АНАЛІЗУ ВЕЛИКИХ ДАНИХ
З ВИКОРИСТАННЯМ МАШИННОГО НАВЧАННЯ**

05.13.05 – комп'ютерні системи та компоненти

Автореферат дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2019

Дисертацією є рукопис

Робота виконана в Харківському національному університеті радіоелектроніки,
Міністерство освіти і науки

Науковий керівник: доктор технічних наук, професор
Хаханов Володимир Іванович, Харківський
національний університет радіоелектроніки,
головний науковий співробітник кафедри
автоматизації проектування обчислювальної
техніки.

Офіційні опоненти: доктор технічних наук, професор
Мірошник Марина Анатоліївна, Український
державний університет залізничного транспорту
МОН України, професор кафедри спеціалізованих
комп'ютерних систем;

доктор технічних наук, професор
Хажмурадов Манап Ахмадович, Національний
науковий центр "Харківський фізико-технічний
інститут" НАН України, начальник відділу
математичного моделювання та дослідження
ядерно-фізичних процесів і систем.

Захист відбудеться "26" вересня 2019 р. о 13-00 годині на засіданні спеціалізованої вченої ради Д64.052.01 в Харківському національному університеті радіоелектроніки за адресою: 61166, місто Харків, пр. Науки, 14.

З дисертацією можна ознайомитись в бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, місто Харків, пр. Науки, 14.

Автореферат розісланий "29" липня 2019 року.

Вчений секретар
спеціалізованої вченої ради

Є.І. Литвинова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми дослідження. Запропоноване дослідження вирішує науково-практичну задачу – введення в інфраструктуру комп'ютерного простору програмної надлишковості у формі моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування. Тема дисертаційної роботи націлена на розробку таких моделей, методів і програмних додатків. Істотний внесок в наукові дослідження, що стосуються інформаційної безпеки та захисту від комп'ютерних загроз, внесли вчені: J. von Neumann, L. Penrose, F. Cohen, D.M. Chase, S.R. White, J.O. Kephart, F. Leitold, L. Adleman, B. Schneier, E. Chien, J. Bruce, J. Bates, R. Greenberg, R. Kusnierz, J. Hruska, F. Skulason, J. Norman, E. Spafford, E. Wilding, V. Bontchev, D. Ferbrache, S. Oxley, J. Norstad, S. Emery, M. Smith, K. van Wyk, S. Staniford, Liang-Jie Zhang, Jia Zhang, Hong Cai, Є. Касперський, С. Новиков, І.Д. Горбенко, В.С. Харченко

Зв'язок роботи з науковими програмами та темами. Розробка теми дисертації здійснювалася відповідно до планів держбюджетних науково-дослідних робіт і міжнародних договорів, виконуваних на кафедрі Автоматизації проектування обчислювальної техніки ХНУРЕ в період з 2007 року, у тому числі: 1) Прикладна держбюджетна НДР № 216 «Енергозберігаючі інформаційні технології на основі паралельних обчислювальних процесів, безпровідних систем і мереж», 2007-2008, № ДР 0107U001540. 2) Договір про дружбу і співробітництво між ХНУРЕ та компанією «Aldec Inc.» (USA) № 04 від 01.11.2011. 3) Фундаментальна держбюджетна НДР № 232 «Теорія й проектування енергозберігаючих цифрових обчислювальних систем на кристалах, що моделюють і підсилюють функціональні можливості людини, 2009-2011, № ДР 0109U001646. 4) Фундаментальна держбюджетна НДР № 269 «Мультипроцесорна система пошуку, розпізнавання та прийняття рішень для інформаційної комп'ютерної екосистеми», 2011-2013, № ДР 0111U002956. 5) Фундаментальна держбюджетна НДР № 258 «Персональний віртуальний кіберкомп'ютер та інфраструктура аналізу кіберпростору», 2012-2014, № ДР 0112U000209. 6) Фундаментальна держбюджетна НДР № 297 «Кіберфізична система – «Розумне хмарне управління транспортом» (Cyber Physical System – Smart Cloud Traffic Control)», 2015-2017, № ДР 0115U-000712 від 04.03.2015. 7) Фундаментальна держбюджетна НДР № 316 "Cyber Physical System – Smart Cyber University", 2017-2019, № ДР 0117U0002524. 8) Проект SEIDA BAITSE "Baltic Academic IT Security Exchange", Blekinge Institute of Technology, Sweden; 2011-2014. 9) Проект 530785-TEMPUS-1-2012-1-PL-TEMPUS-JPCR «Curricula Development for New Specialization: Master of Engineering in Mi-

crosystems Design (MastMEMS)» сумісно з університетом «Львівська політехніка», Київським національним університетом, Технічним університетом м. Лодзь (Польща), Ліонським університетом (Франція), Університетом м. Ільменау (Німеччина), Університетом м. Павія (Італія), 2012-2016. 10) Проект 544455-TEMPUS-1-2013-1-SE-TEMPUS-JPCR «Educating the Next Generation Experts in Cyber Security: the new EU-recognized Master's program (ENGENSEC)», 01 Dec 2013 – 30 Nov 2017.

Автор дисертаційної роботи брав участь у виконанні зазначених договорів і програм як розробник, менеджер і програміст кіберфізичної інфраструктури захисту кіберпростору у вигляді програмних засобів перевірки, діагностування шкідливих програм і атак, що дає можливість виконувати їх моделювання із залученням існуючих додатків і malware бібліотек.

Науково-практична задача полягає у введенні в інфраструктуру комп'ютерного простору програмної надмірності у формі моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування.

Сутність дослідження – розробка моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування за рахунок введення обчислювальної надмірності в інфраструктуру кіберпростору.

Об'єкт дослідження – хмарні і edge-computing технології, засновані на високопродуктивних дата-центрах, комп'ютерних гаджетах, системах і мережах, що виконують сервісні функції зберігання та аналізу великих даних.

Предмет дослідження – структурно-логічні моделі, методи, засоби тестування деструктивних компонентів, захисту індивідуального та колективного сервіс-комп'ютерингу від кіберзагроз.

Мега дослідження – істотне скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування за рахунок введення обчислювальної надмірності в інфраструктуру кіберпростору.

Задачі дослідження:

1) Удосконалити структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів на основі використання дедуктивного аналізу обчислювальних систем.

2) Розробити сигнатурно-кубітні методи синтезу еталонних логічних схем malware-функціональностей і паралельного моделювання malware-driven великих даних для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці.

3) Розробити сигнатурно-кубітну модель активного online cyber security комп'ютингу для моніторингу вхідних потоків malware-даних і управління процесом видалення деструктивних компонентів.

4) Удосконалити засоби захисту кіберпростору шляхом логічного тестування і діагностування атак і шкідливих компонентів на основі використання алгоритмів машинного навчання.

5) Розробити метод атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів і метод перевірки поліморфних шкідливих програм на основі врахування контрольних сум Portable Executable секцій у виконуваних файлах і застосування апарату інтелектуального аналізу даних.

6) Виконати тестування і верифікацію розроблених програмних засобів тестування, перевірки та діагностування шкідливих програм шляхом емуляції атак на основі існуючих malware бібліотек.

Наукова новизна результатів дисертаційної роботи:

1) *Удосконалено* структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware.

2) *Вперше запропоновано* методи синтезу еталонних логічних схем malware-функціональностей, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці.

3) *Вперше розроблено* модель активного online cyber security комп'ютингу, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкодію процесів моніторингу вхідних потоків malware-даних і управління видаленням деструктивних компонентів.

4) *Вперше запропоновано* метод атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів і метод перевірки поліморфних шкідливих програм на основі врахування контрольних сум Portable Executable секцій у виконуваних файлах і застосування апарату інтелектуального аналізу даних.

5) *Удосконалено* засоби захисту кіберпростору, які відрізняються використанням моделей і методів сигнатурно-логічного тестування атак, пошуку криптопримітивів у троянських програмах-шифрувальниках на основі використання алгоритмів машинного навчання, що дає можливість істотно скоротити час відновлення працездатності обчислювальної структури.

Практичне значення одержаних результатів досліджень полягає у

– тестуванні, верифікації і впровадженні розроблених програмних засобів перевірки, діагностування шкідливих програм і атак, що дає можливість виконувати їх моделювання із залученням існуючих додатків і malware бібліотек;

– програмній реалізації методу атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів, який відрізняється застосуванням апарату інтелектуального аналізу даних, що дає можливість визначати вірогідну оцінку небезпеки URL-адреси на основі його атрибутів;

– програмній реалізації методу перевірки поліморфних шкідливих програм, який відрізняється інваріантністю до детермінізму сигнатур в код і урахуванням тільки контрольних сум Portable Executable (PE) секцій у виконуваних файлах, що дає можливість поліпшити продуктивність процедур діагностування деструктивних компонентів.

Обґрунтованість наукових положень. Отримані в процесі виконання досліджень наукові висновки і практичні результати є достовірними, що підтверджується точністю детектування і класифікацією прикладів загроз нульового дня, серед яких нові версії кріптолокерів і складних загроз (Advanced Persistent Threats – АРТ); позитивними відгуками вчених і фахівців на доповіді на міжнародних конференціях, присвячених кібербезпеці, таких як Virus Bulletin 2015 Празі, Чехія, 2018 у Монреалі, Канада, OpenStack Summit 2015 у Ванкувері, Канада, і OpenStack Summit 2015 в Остіні, штат Техас, США, IEEE East - West Design & Test Symposium (EWDTS'2017) 2017 в Novy Sad, Serbia.

Впровадження результатів дисертації. Результати дисертації у складі моделей, методів та інфраструктури впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акти про впровадження від 20.05.2019, 21.05.2019); у науково-виробничу діяльність компанії Design & Test Lab (довідка від 18.05.2019), у навчальний процес Blekinge Institute of Technology (ВТН), Karlskrona, Sweden (лист ‘Statment of Reseach Results Impact on University Education Program’ від 29.05.2019).

Особистий внесок здобувача. Всі наукові і практичні результати отримані автором особисто. У роботах, опублікованих зі співавторами, здобувачеві належать: [1] – мультипроцесорна архітектура для обробки великих даних; [2] – структурна модель malware-аналізу; [3] – модель аналізу інформаційної безпеки кіберпростору; [4] – технології кіберзахисту для корпоративних мереж; [5] – інфраструктура кіберпростору для формування інформаційної безпеки; [6] – удосконалені структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware; [7] – методи синтезу еталонних логічних схем malware-функціональностей для паралельного моделювання malware-driven великих даних; модель активного online cyber security комп'ютингу; [8] – метод визначення поліморфного шпигуна; [9] – метод виявлення URL-адрес за допомогою частотних шаблонів; [10] – аналітична модель оцінки збитків підприємства від шкідливих програм; [11] – модель функціональної перевірки системи на кристалі; [12] – опис програмних засобів для реалізації моделей системного рівня; [13] – модель аналізатора вихідного

коду для мов опису апаратури; [14] – метод інтелектуального аналізу даних для функціональної перевірки системи на кристалі; [15] – технологія інтелектуального аналізу даних для функціональної перевірки SoC; [16] – опис та тестування програмних модулів для виявлення троянських включень у програмному забезпеченні; [17] – метод детектування апаратних троянських включень; [18] – методи діагностування і структури даних для інформаційного захисту; [19] – модель захисту індивідуального кіберпростору; [20] – аналіз і тестування програмних модулів; [21] – структурна модель для отримання інформації у великих даних; [22] – модель для захисту кіберпростору в хмарі; [23] – структурна модель криптоперетворення та тестування програмних модулів для операційної системи Android; [24] – метод виявлення таргетованих атак у хмарі; [25] – архітектура безпеки з відкритим кодом для захисту від цільових атак; [26] – опис програмних засобів для реалізації моделі реагування на хмарні інциденти; [27] – тестування програмних модулів для реалізації методу пом'якшення; [28] – аналіз кібератак; [29] – застосування штучного інтелекту для криптоаналізу; [30] – моделі дескрипторів; [31] – модель пошуку шкідливого коду у програмних продуктах.

Апробація результатів дисертації. Результати роботи були представлені та обговорені на наступних конференціях: IEEE East-West Design and Test Symposium 2007, 2016 (Yerevan, Armenia), 2009 (Moscow, Russia), 2010 (Saint Petersburg, Russia), 2011 (Sevastopol, Ukraine), 2014 (Kyiv, Ukraine), 2015 (Batumi, Georgia), 2017 (Novi Sad, Serbia); International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 2008 (Lviv-Polyana, Ukraine); Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті», 2012, (Харків, Україна); the 13th International Conference «The Experience of Designing and Application of CAD Systems in Microelectronics», CADSM 2007, 2009, 2015 (Lviv, Ukraine); the Virus Bulletin International Conference, 2015 (Prague, Czech Republic), 2017 (Madrid, Spain), 2018 (Montreal, Canada); OpenStack Summit, 2015 (Vancouver, Canada), 2016 (Austin, TX, USA); Cyber Security Conference UISGCON14, 2018 (Kyiv, Ukraine). Автор також брав участь у конкурсах інноваційних проєктів та розробок, як запрошений експерт, спікер, член програмного комітету та журі, серед яких Міжнародна студентська конференція і конкурс наукових робіт з питань інформаційної безпеки «CyberSecurity for the Next Generation», 2011-2014, Kaspersky Lab.

Публікації. Результати дисертаційної роботи відображені в 31 друкованій праці: 3 розділи у закордонних монографіях (з них 1 входить до наукометричної бази Scopus), 7 статей (з них 5 – у наукових журналах, включених до «Переліку наукових фахових видань України»; 2 статті в міжнародних наукових журналах за кордоном; 4 статті входять до міжнародних наукометричних баз),

а також у 21 міжнародній науковій конференції (з них 13 за кордоном, 12 входять до наукометричної бази Scopus). Здобувач має 13 публікацій у наукометричній базі Scopus та індекс Хірша $h=3$.

Дисертаційна робота має 241 сторінку (з них 218 представляють основний текст) і містить: 5 розділів, 48 рисунків, 14 таблиць, список джерел з 160 назв (на 16 с.), 4 додатки (на 16 с.), анотації на 27 с.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обгрунтовано актуальність завдань, які вирішуються в дисертаційній роботі, сформульована мета дослідження, а також викладені наукова новизна і практична цінність отриманих результатів.

У **першому розділі** наводиться аналітичний огляд існуючих моделей, методів і технологій захисту індивідуального сервіс-комп'ютингу. Визначаються переваги і недоліки найбільш затребуваних моделей і методів, опублікованих в спеціальній літературі: матеріалах конференцій і наукових журналах. На основі проведеного аналізу сформульовано мету і задачі дослідження, орієнтовані на усунення проблемних місць і недоліків існуючих моделей і методів у контексті їх реалізації в інфраструктурі захисту індивідуального сервіс-комп'ютингу.

Функція мети (Z) – мінімізація проміжку часу між моментом запуску атаки (A) на кіберпростір і моментом її діагностування (D), протягом якого обчислювальний сервіс залишається скомпрометованим (C), що одночасно дозволяє поліпшити якість сервісу шляхом забезпечення доступності, цілісності і конфіденційності оброблюваної інформації на період атаки; мінімізації витрат на відновлення працездатності сервісу і фінансових втрат від його простою (TDT) за рахунок введення мінімально необхідної надмірності в інфраструктуру діагностування (I):

$$Z = F(TDT, TC, I) = \min[\frac{1}{3}(TDT + TC + I)],$$

де TC – час, на протязі якого обчислювальний сервіс залишається скомпрометованим з моменту запуску атаки зловмисником,

$$TC = t(D) - t(A),$$

де $t(D)$ – момент детектування атаки, $t(A)$ – момент запуску атаки.

У **другому розділі** наводяться *удосконалені* структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware. Пропонується блокчейн-технологія і математичний апарат створення інфраструктури програмно-апаратних телекомунікаційних інформаційних кіберфізичних систем (КС), орієнтована на захист від несанкціонова-

ного доступу до сервісів, визначених у специфікації системи, шляхом проникнення через легальні інтерфейси взаємодії компонентів, що мають уразливість. Інфраструктура захисних сервісів створюється разом з кіберсистемою і супроводжує останню протягом всього життєвого циклу, обслуговуючи всі наступні модифікації КС, і сама постійно підвищує свій інтелект шляхом поповнення історії та бібліотек конструктивних і деструктивних компонентів. Функція мети представлена підвищенням ефективності сервісного обслуговування на основі стандартів тестування, граничного сканування і спеціальних технологій діагностування та відновлення невразливості КС, яка визначається мінімальним значенням рівня вразливості, часу відновлення працездатності Т і нефункціональної програмно-апаратної надмірності Н:

$$E = F(L, T, H) = \min\left[\frac{1}{3}(L + T + H)\right],$$

$$Y = (1 - P)^n;$$

$$L = 1 - Y^{(1-k)} = 1 - (1 - P)^{n(1-k)};$$

$$T = \frac{(1-k) \times H^s}{H^s + H^a}; \quad H = \frac{H^a}{H^s + H^a},$$

де L – доповнення до рівня невразливості Y, яке залежить від тестопридатності КС k, ймовірності P існування вразливостей і числа невиявлених деструктивних n. Час тестування і діагностування залежить від тестопридатності архітектури k, помноженої на число структурних компонентів інфраструктури, віднесених до загальної кількості елементів КС. Надмірність залежить від структурної складності тестопридатності надбудови, поділеної на програмно-апаратну складність КС. Надмірність інфраструктури забезпечує задану глибину діагностування вразливостей за час, що визначається замовником.

Пропонується математичний апарат інфраструктури захисного сервісу, що містить метрику, алгебру, структури даних і моделі оцінювання якості взаємодії процесів, явищ, об'єктів і компонентів у кіберпросторі і кіберсистемі, необхідних при створенні ефективних двигунів для обчислювальних процедур аналізу даних в процесах тестування проникнень і відновлення невразливості. Вводиться модифікована модель критерію скалярної і векторної якості оцінювання бінарних відношень, яка відрізняється використанням функції неналежності та кодової відстані Хеммінга, що забезпечує лінійність зміни чисельного значення критерію від 0 до 1 в міру збільшення «відстані» від повного збігу двох об'єктів до максимально можливого, коли кодова відстань дорівнює $d(m, A) = k$. Критерій може бути використаний при оцінюванні взаємодії об'єктів у реальному масштабі часу в задачах тестування, діагностування функціональних порушень, вразливостей. Для синтезу тестів застосовується апарат булевих похідних, що призначений для перевірки суттєвості змінних і компоне-

нтів КС, включаючи аналіз суттєвості деструктивності (уразливості і проникнення) для стану кіберсистеми. Запропонована процес-модель синтезу тестів для тестування і діагностування вразливостей може бути використана як вбудований компонент інфраструктури сервісного обслуговування КС.

Пропонується дедуктивний метод пошуку вразливостей в КС, основна ідея якого полягає в аналізі зіставлення вхідних і вихідних даних кіберсистеми з метою виявити деструктивні проникнення або уразливості шляхом виконання процедур порівняння між штатними (функціональними) режимами і ситуаціями, що викликають підозру. Для імплементації методу в інфраструктуру захисних сервісів необхідно мати графову модель логіки функціонування кіберсистеми, яка досить просто може бути трансформована до системи логічних рівнянь, придатної для дедуктивного аналізу. Пропонується модель дедуктивно-паралельного синхронного аналізу вразливостей (проникнень) кіберсистеми (об'єкта), яка дозволяє за одну ітерацію обробки структури обчислити всі деструктивні компоненти, що перевіряються на тест-векторі. Мета дедуктивного аналізу – визначити якість синтезованого тесту щодо повноти покриття ним вразливостей, а також побудувати таблицю перевірки тестовими наборами усіх виявлених вразливостей КС для виконання процедур діагностування. Така модель заснована на розв'язанні рівняння $L = T \oplus F$, де $F = (F_{m+1}, F_{m+2}, \dots, F_i, \dots, F_n)$, $i = m+1, \dots, n$ – сукупність функцій справної (коректної) поведінки КС; m – число його входів; $Y_i = F_i(X_{i1}, \dots, X_{ij}, \dots, X_{in_i})$ – n_i -входовий i -й елемент схеми, що реалізує F_i для визначення стану лінії (виходу) Y_i на тест-векторі T_t ; тут X_{ij} – j -й вхід i -го елемента; тест $T = (T_1, T_2, \dots, T_t, \dots, T_k)$ – упорядкована сукупність двійкових векторів, визначена в процесі справного моделювання на множині вхідних, внутрішніх і вихідних ліній, об'єднана в матрицю $T = [T_{ij}]$.

Спільна апаратурна реалізація ДФ для двохвходових елементів And , Or на вичерпному тесті представлена універсальним функціональним примітивом (рис. 1) дедуктивно-паралельного аналізу несправностей. У симуляторі представлені булеві (x_1, x_2) і реєстрові (X_1, X_2) для кодування вразливостей входи, змінна вибору типу справної функції (AND, OR), вихідна реєстрова змінна Y . Стани двійкових входів x_1, x_2 і змінна вибору елемента визначають одну з чотирьох дедуктивних функцій для отримання вектора Y перевірки несправностей.

Застосування такого симулятора дає можливість трансформувати функціональну модель F коректної поведінки КС в дедуктивну L, яка інваріантна в сенсі універсальності тестовим наборам і не передбачає в процесі моделювання використовувати модель F. Тому симулятор, як апаратна модель ДФ, є ефективним двигуном дедуктивно-паралельного моделювання КС, що підвищує швидкість аналізу кіберсистем в 10 – 1000 разів у порівнянні з програмною реалізацією. Але при цьому співвідношення обсягів моделей коректного

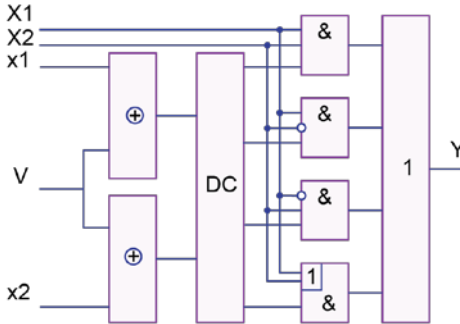


Рис. 1. Симулятор несправних примитивів

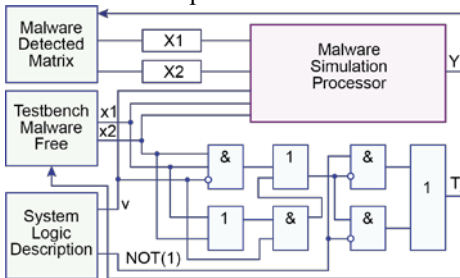


Рис. 2. HFS-структура апаратного моделювання

моделювання та аналізу вразливостей становить 1:10. Підхід апаратного аналізу деструктивний, спрямований на розширення функціональних можливостей вбудованих засобів моделювання, які можна зберігати на хмарі і постійно ними користуватися для верифікації інфраструктури захисту КС. Обчислювальна складність обробки проекту $Q=(2n2r)/W$, де r – час виконання регістрової операції (And, Or, Not); W – розрядність регістра.

Для апаратної реалізації дедуктивно-паралельного моделювання на основі запропонованого симулятора може бути використана обчислювальна структура, представлена на рис. 2. Особливість схемної реалізації полягає в спільному виконанні двох операцій: однобітової – для емуляції функцій логічних елементів And, Or і паралельної – для обробки багаторозрядних векторів несправностей шляхом виконання операцій логічного множення, заперечення і складання.

Функціональне призначення основних блоків (пам'ять і процесор): 1. $M=[M_{ij}]$ – квадратична матриця моделювання деструктивних проникнень (ДП), де $i, j = 1, q$; q – загальне число ліній в оброблюваній КС. 2. Вектори збереження станів коректного моделювання, визначені в моменти часу $t-1$ і t , необхідні для формування дедуктивних функцій примитивів. 3. Модуль пам'яті для зберігання опису КС у вигляді структури логічних елементів. 4. Буферні

регістри, розмірністю q , для зберігання операндів і виконання регістрових паралельних операцій над векторами ДП, що зчитані з матриці M . 5. Блок коректного моделювання для визначення двійкового стану виходу чергового оброблюваного логічного елемента. 6. Дедуктивно-паралельний симулятор, який обробляє за один такт дві реєстрових змінних $X1, X2$ з метою визначення вектора ДП, що транспортуються на вихід логічного елемента Y .

Перевага запропонованої структури моделювання ДП. 1. Суттєве зменшення кількості модельованих ДП, які визначаються тільки числом збіжних розгалужень, що становить до 20% від загального числа ліній. 2. Зниження обсягу пам'яті, необхідного для зберігання матриці модельованих ДП. 3. Простота реалізації Hardware Vulnerability Simulator (HVS) в апаратному виконанні, що дозволяє на порядок збільшити швидкість моделювання ДП. 4. Використання HVS як першої фази дедуктивно-топологічного методу, який ґрунтується на результаті обробки розгалужень, що сходяться, для швидкодійного аналізу деревоподібних структур.

В третьому розділі пропонуються *нові методи* синтезу еталонних логічних схем malware-функціональностей, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці. Вводиться нова модель активного online cyber security комп'ютигу, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкість процесів моніторингу вхідних потоків malware-даних і управління видаленням деструктивних компонентів. Пропонуються унітарно кодовані кубітно-матричні моделі, структури даних, обчислювальні архітектури і методи паралельного логічного аналізу деструктивних кодів у кіберфізичному просторі. Вводяться кубітні векторні структури даних для опису параметрів змінних, що беруть участь у формуванні еталонних зразків (паттернів) деструктивних вихідних кодів. Вводиться сигнатурно-кубітний процесор активного online кіберфізичного cyber security комп'ютигу (CSC) на основі моніторингу вхідних malware-даних і їх моделювання на еталонних логічних схемах malware-функціональностей з метою подальшого актуаторного управління процесом видалення деструктивних компонентів.

Cyber Security Computing (CSC) – галузь знань, що займається розвитком теорії і практики надійного метричного online управління кіберзахисту великих даних, віртуальних, фізичних (природних) і соціальних процесів і явищ в комп'ютерних дата-центрах і мережах на основі точного цифрового моніторингу деструктивних компонентів кіберфізичного простору за допомогою інтелектуальних пошуково-аналітичних сервісів і розумних сенсорів. CSC – процес тестування, моніторингу, діагностування та активації сигналів деструкції

шкідливих компонентів на основі метричних відносин між malware і software в кіберфізичному просторі.

CSC-процес – спостережувана взаємодія механізмів malware і software в часі і просторі на основі моніторингу та актуації метричних відносин для досягнення мети у вигляді усунення malware при виділених ресурсах.

Malware-функціональність (MF) являє собою структуру взаємопов'язаних логічних елементів, яка забезпечує цифрову реалізацію деструктивної поведінки об'єкта в заданому просторі software змінних.

Malware-змінна (MV) визначається упорядкованим універсумом примітивних значень, який формує проекцію поведінки об'єкта на векторі змінних, що створює malware-функціональність.

Логічний malware-елемент (ML) являє собою еталонне відображення значень багатозначної змінної в двійковий кубітний вектор, заданий на упорядкованому універсумі примітивних значень.

Значення (LV) змінної – унікальна примітивна властивість об'єкта, що має пустий перетин з іншими примітивами, яке в суперпозиції з ними складає універсум.

Таким чином, проглядається структурована ієрархія введених понять (рис. 3): (CSC – MF – MV – ML – LV), яка формує можливі архітектурні рішення malware-комп'ютингу.

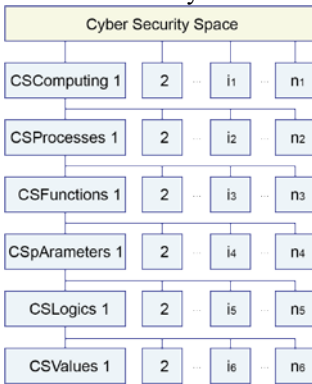


Рис. 3. Ієрархія malware-комп'ютингу

$$W = (U, Q, L),$$

$$U = (U_1, U_2, \dots, U_i, \dots, U_n);$$

$$\bigcup_{i=1}^n U_i = U; U_i \cap U_k = \emptyset; \quad i \neq k$$

$$Q = (Q_1, Q_2, \dots, Q_i, \dots, Q_n);$$

$$\bigcup_{i=1}^n Q_i = Q; Q_i \cap Q_k = \emptyset; \quad i \neq k$$

$$Q_i = (Q_{i1}, Q_{i2}, \dots, Q_{ij}, \dots, Q_{im}); Q = [Q_{ij}];$$

$$U_i = (U_{i1}, U_{i2}, \dots, U_{ij}, \dots, U_{im}); U = [U_{ij}];$$

$$L = f[Q] = (Q_1 \circ Q_2 \circ \dots \circ Q_i \circ \dots \circ Q_n)$$

$$\circ = \{\wedge, \vee, \oplus\};$$

$$U_{ij} \in U_i \in U; Q_{ij} \in Q_i \in Q; Q_i \in U_i; Q \in U;$$

$$Q_{ij} = 1 \leftarrow \max \mu(R, U_{ij}).$$

ML-рівень архітектури характеризується синтезом логічної схеми, де кожен елемент має одну багатозначну регістрову змінну, яка фактично представлена кубітним вектором, де число одиничних координат може бути більше одиниці. Дана властивість кубіта дає можливість створювати компактні структури даних malware-функціональностей з метою їх паралельної обробки. Для

виконання квантового методу моделювання на кубітних структурах даних необхідно унітарно закодувати вхідні символічні дані за допомогою таблиць-універсумів значень, що відповідають кожній змінній.

Завдання пов'язані зі створенням моделі, методу і процесора malware-комп'ютингу, спрямованого на автоматичний синтез і аналіз кубітних логічних схем, орієнтованих на моніторинг, моделювання, розпізнавання і деструкцію шкідливих кодів.

Вводиться аналітична модель W кубітно-логічного процесора CSC-комп'ютингу, яка оперує двома матрицями: універсумів U примітивів і кубітних функціональностей Q , а також логічними примітивами L , що інтегрують функціональності в комбінаційну схему CSC-процесора:

Архітектура метричної взаємодії U -матриці універсумів з потоком даних R для обчислення функцій належності $\mu(R, U)$, з метою отримання Q -матриці значень і подальшого L -об'єднання кубітів у комбінаційну схему кіберсоціального процесора, представлена на рис. 4.

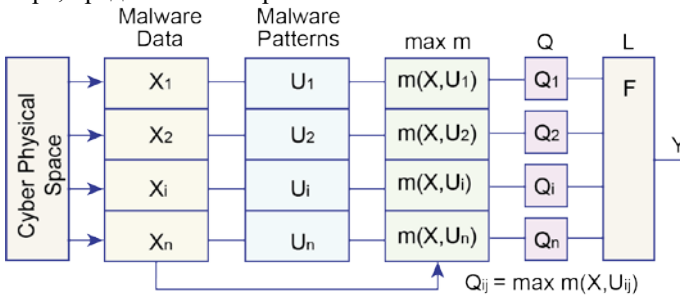


Рис. 4. Архітектура для синтезу CSC-процесора

Тут вхідний потік модельованих великих деструктивних даних R має такий же формат, як U -, Q -матриці і комбінаційна схема процесора. Алгоритм синтезу Q -матриці полягає у визначенні максимального значення функції належності вхідного фрейму розглянутої змінної до одного з значень відповідного рядка матриці універсумів. В результаті такого порівняння за всіма координатами U -матриці формуються поодинокі координати кубітної матриці, де кожен рядок являє собою примітивну функціональність по розглянутій змінній. Разом всі рядки Q -матриці створюють комбінаційну схему логічного CSC-процесора для моделювання будь-якого вхідного впливу з метою визначення його належності до даного еталону деструктивного процесу або явища.

Розглядається процесорна архітектура активного online кіберфізичного cyber security комп'ютингу, яка характеризується моніторингом вхідних потоків malware-даних, їх подальшим моделюванням на еталонних логічних схемах malware-функціональностей, що дає можливість в online режимі актуально управляти процесом видалення деструктивних компонентів.

Логічна структура, представлена на рис. 5, характеризується комп'ютирною архітектурою malware-аналітики, яка має всі вісім компонентів моделі універсального обчислювача. Вони взаємодіють між собою за формулою: вхід R ініціює команди для виконання CSC-процесу, який починається з отримання D-ресурсів: фінансових, кадрових, інформаційних, що надходять на виконавчий механізм E. Він активує сенсори, що передають інформацію по шині моніторингу M для подальшого формування потоку S-даних, призначеного для синтезу U-матриці універсуму змінних для опису CSC-процесу за допомогою універсумів вербальних значень кожної змінної. Остання слугує базовим форматом для синтезу X-матриці векторів вхідних даних по кожній змінній для виконання ітерації моделювання відносно Q-матриці, що задає сукупність кубітних двійкових векторів для опису еталонного malware-патерна за всіма змінними. Це дає можливість на основі &-операції обчислювати функцію належності μ , що має вихід візуалізації стану V CSC-процесу, і відповідну Y, у вигляді чисельного значення кількості різних двійкових однойменних координат, отриманого при порівнянні матриць X і Q, який визначається за допомогою паралельної операції $(X \wedge \text{not}Q) = Y$. Вона дає можливість синтезувати актуаторні вербальні сигнали W, відповідні одиничним значенням координат вихідної матриці Y в форматі U-матриці, які по шині A ініціюють обчислювальні процедури, спрямовані на усунення розходжень між X-матрицею і еталонною Q-матрицею malware-паттерну шляхом корекції координат X-матриці за допомогою інфраструктури E, що забезпечує виконання CSC-процесу компанії для отримання сервісів або продукції P.

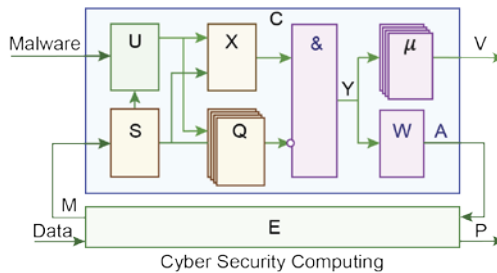


Рис. 5. Логічна матрична паралельна схема CSC-комп'ютирну

Архітектура CSC-комп'ютирну характеризується кінцевою множиною Q-матриць для паралельної обробки вхідних X-матриць з метою отримання μ -матриць, також обчислюються паралельно, що представляє собою універсум функцій приладдя чергового вхідного впливу malware-паттернам, серед яких вибирається мінімальна оцінка відмінностей і скалярні оцінки для усунення malware.

Оригінальність моделі кубітно-матричного процесора, що відповідає архітектурі, представленої на рис. 5, характеризується використанням кубітних матриць даних для виконання операцій CSC-комп'ютингу, структура якого містить механізми управління С, і виконання Е, з'єднані шинами моніторингу М і актуації А. Тут блок управління має зовнішній вхід команд R і вихід візуалізації стану CSC-процесу V, а блок виконання містить зовнішній вхід даних D і вихід сервісів або продукції P, що в сукупності становить вісім компонентів, які характеризують CSC-комп'ютинг. Він починається з визначення вхідного потоку даних S, що надходить від сенсорів виконавчого механізму, обслуговуючого CSC-процес, який призначений для синтезу U-матриці універсуму змінних і опису CSC-процесу за допомогою універсумів вербальних значень кожної змінної, яка є базовим шаблоном для синтезу X-матриці вхідних даних і наступного моделювання сумісно з вже визначеними Q-матрицями, що задають множину еталонних malware-патернів. Це дає можливість на основі &-операції паралельно обчислювати n матриць $Y_i = (X \wedge \text{not} Q_i)$, $i=1, \dots, n$, моделювання, аналіз яких на мінімальне число одиничних координат $\min(Y_i=1)$, $i=1, \dots, n$, дозволяє визначити номер i матриці, що має мінімальне значення з n функцій належностей $\mu = \min_i \mu_i \leftarrow \mu_i = \sum_{j=1, k}^{r=1, m} Y_{ijr}$, які формують чисельні значення відмінностей двійкових однойменних координат, отриманих при порівнянні матриці X і n матриць Q. Це дає можливість синтезувати матрицю актуаторних вербальних сигналів W, відповідну одиничним значенням координат вихідної матриці Y, що має $\min \mu$, в форматі вербальних значень U-матриці, які ініціюють обчислювальні процедури у виконавчому механізмі Е, спрямовані на усунення розходжень між X-матрицею і Q-матрицею malware-паттерну з $\min \mu$, шляхом корекції координат X-матриці, що забезпечує оптимальне виконання мети P CSC-процесу.

Таким чином, запропоновано логічний процесор для паралельного моделювання та розпізнавання malware-паттернів у потоках великих даних на основі створення інтерпретативних кубітних матричних моделей, методів і архітектур CSC-комп'ютингу, спрямованого на автоматичний синтез і аналіз логічних схем, орієнтованих на моніторинг і управління CSC-процесами та явищами для усунення деструктивних компонентів у кіберпросторі.

У **четвертому розділі** розглядаються модель загроз кіберпростору, а також методи діагностування кібератак на кіберпростір з використанням алгоритмів машинного навчання на основі великих даних, що дозволяють виявити загрози, запропоновані у моделі загроз кіберпростору. Пропонуються методи виявлення кіберзагроз, які здатні навчатися на великих даних, з метою виявлення кібератак, що можуть бути реалізовані у вигляді Інтернет посиланнях, поліморфних шкідливих програмах (polymorphic malware) та троянських програ-

мах-шифрувальниках, що займаються здирництвом: 1) Метод атрибутно-орієнтованого впізнання Інтернет посилань з використанням частотних шаблонів, що може провести оцінку атрибутів та відрізнити доброякісні, фішинг та шкідливі посилання. 2) Метод детектування поліморфних шкідливих програм, що дозволяє виявляти поліморфні шкідливі програми за допомогою аналізу хешів PE секцій та пошуку схожих секцій у бібліотеці шкідливого коду. Після підтвердження детектування PE файлу нові хеші секцій додаються до бібліотеки. 3) Метод пошуку криптопримитивів у троянських програмах-шифрувальниках. Завдяки цьому методу можливо відповісти, які алгоритми шифрування використовував здирник та чи можливо дешифрувати зашифровані файли користувача або організації.

У **п'ятому розділі** у межах практичної реалізації виконана розробка і тестування компонентів інфраструктури Cyber Security. Пропонується хмарний сервіс для виявлення кібератак на основі аналізу великих даних, який включає три основних компоненти: 1) Сервер, керуючий потоком вхідних і вихідних даних про кібератаки. 2) Мультисканер з препроцесором для статичного аналізу. 3) Пісочниця (Sandbox), яка призначена для автоматизованого запуску шкідливого коду з метою проведення динамічного аналізу. Для реалізації сервісу створена інфраструктура на основі гіпервизора VMWare ESXi (рис. 6).

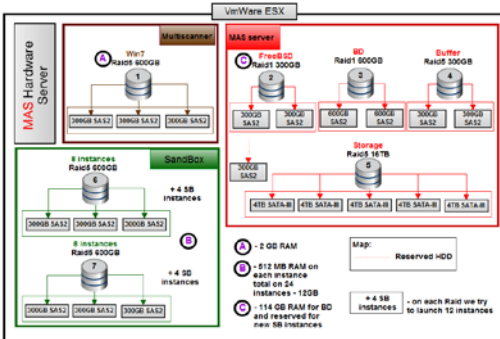


Рис. 6. Інфраструктура VMWare ESXi

шкідливі відповідно в тому ж наборі. Середнє число помилкових спрацьовувань методу FPA вище, ніж IDS і GSB сервісів, і складає 0,04% (9 посилань) від загального числа. В контексті детектування підозрілих Інтернет посилань даний показник не є критичним, як, наприклад, для додатків, оскільки користувач може зайти на заблоковану сторінку, проігнорувавши попередження системи безпеки. При цьому середня кількість пропущених шкідливих і фішинг посилань вища у систем IDS і GSB і становить 5,32 і 5,06% відповідно при тому, що FPA метод пропустив всього 0,59% небезпечних посилань. Даний

За результатами детектування Інтернет посилань на основі аналізу атрибутів з використанням частотних паттернів можна зробити такий *висновок*: метод дозволяє визначити 4,93% посилань як шкідливі на підставі аналізу їх атрибутів із загального числа посилань "in-the-wild". IDS система і GSB сервіс зміг визначити 0,16 і 0,43% посилань як

параметр є найбільш істотним в системах блокування шкідливих і фішинг сайтів, оскільки дозволяє попередити користувача про потенційну небезпеку при відвідуванні зараженого веб-ресурсу.

В результаті порівняльного аналізу методу сигнатурного детектування IDS системи Suricata і методу детектування з використанням чорних списків Google Safe Browsing з розробленим методом Frequency Patterns Analysis (FPA) на основі аналізу атрибутів Інтернет посилань ми отримали значну перевагу методу за кількістю детектируваних шкідливих і фішинг посилань. При цьому розроблений метод FPA має прийнятний відсоток помилкових спрацьовувань (0,04%), що є незначним недоліком методу на даний момент.

Висновок за результатами детектування поліморфних шкідливих програм. Як було показано вище, технологія поліморфізму може значно захистити нові частини програм-шпигунів від виявлення антивірусом нульового дня, роблячи себе практично невидимим в комп'ютерній системі. Більш того, після установки поліморфні бекдори можуть запускати процедуру поновлення для завантаження нової версії шпигунського ПЗ, що збільшує термін його служби на зараженому комп'ютері. Запропоновано також спосіб виявлення поліморфних шпигунських програм і те, як цей підхід заснований на динамічному аналізі зразків. Після запуску поліморфна програма-шпигун виявляє своє шкідливе корисне навантаження безпосередньо в пам'яті процесу. Цей метод може бути успішно виявлений за допомогою правил Yara, спеціально створених для родин Nrgbot і Shiz. Беручи цю інформацію до уваги, можна запропонувати описаний метод виявлення численних інфекцій, наприклад, в корпоративній мережі. Використовуючи описані методи, адміністратор або інженер з безпеки може створити правило Yara для певного сімейства шпигунських програм і почати виявляти активне зараження в мережі. Як тільки вірус виявлений, керівництво з видалення може допомогти вилікувати систему.

Висновок за результатами пошуку криптопримітивів. Результати експериментів підтвердили первинну гіпотезу про можливість ідентифікації криптопримітивів в упакованому і деобфускрованому коді вимагачів. Випадки, представлені в табл. 1-5, мають правильно розпізнаний фрагмент коду з мінімальним відстанню Левенштейна.

Таблиця 1. Зразки здирників, використані в експериментах

Ransomware	SHA256
TeslaCrypt	9e3827dff24d1da72cb3d423bddf4cd535fa636062e4ea63421ef327fec56ad
GlobeIm- poster	a0e5bced56025f875721043df981c400fc28e4efc68ffe42ac665633de085ab1
MoneroPay	ababb37a65af7c8bde0167df101812ca96275c8bc367ee194c61ef3715228ddc

Таблиця 2. Виявлення AES (ключова експансія) в TeslaCrypt

Експеримент №	1	2	3	4	5
Очікуване розташування	0	1500	3000	10000	20000
Відповідне місцеположення	115	1473	2986	10006	19953
Відстань Левенштайна	95	60	93	76	75
Оцінка	FALSE	TRUE	FALSE	FALSE	FALSE

Таблиця 3. Виявлення AES (розширення ключа) в GlobeImposter

Експеримент №	1	2	3	4	5
Очікуване розташування	100	1000	4400	10000	20000
Відповідне місцеположення	399	999	4425	9968	19991
Відстань Левенштайна	61	113	50	132	91
Оцінка	FALSE	FALSE	TRUE	FALSE	FALSE

Таблиця 4. Виявлення RC4 (PRGA) у GlobeImposter

Експеримент №	1	2	3	4	5
Очікуване розташування	0	500	800	1000	1500
Відповідне місцеположення	340	340	828	1063	1553
Відстань Левенштайна	20	20	76	75	83
Оцінка	TRUE	TRUE	FALSE	FALSE	FALSE

Таблиця 5. Виявлення Salsa20 (quaterround) у MoneroPay

Експеримент №	1	2	3	4	5
Очікуване розташування	0	100	1000	1500	3000
Відповідне місцеположення	2	100	1000	1500	3094
Відстань Левенштайна	118	146	177	619	389
Оцінка	TRUE	FALSE	FALSE	FALSE	FALSE

Таким чином, показана можливість ефективного розпізнавання криптопримітивів у кодї здирників на основі коду асемблера ethalon з симетричних шифрів, які розглядаються як шаблони пошуку. Запропонований метод успішно виявив криптографічний код симетричних шифрів AES, RC4, Salsa20 в здирників TeslaCrypt, GlobeImposter і MoneroPay. Проте заснований на сигнатурах підхід з використанням правил Yaga та відкритим вихідним кодом і KANAL в PEiD, показав низькі можливості виявлення. Описаний спосіб може бути додатково поліпшений, щоб автоматично генерувати правило Yaga для виявлення фрагмента коду, що збігся, вимагача, відповідального за симетричне шифрування даних.

ВИСНОВКИ

Проведене дослідження вирішує науково-практичну задачу – введення в інфраструктуру комп'ютерного простору програмної надмірності в формі моделей, методів і програмних додатків для істотного скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування.

Для досягнення поставленої мети в роботі були вирішені завдання, які дозволили отримати результати, що мають наукову новизну:

1) *Удосконалені* структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware.

2) *Нові* методи синтезу еталонних логічних схем malware-функціональних, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці.

3) *Нова* модель активного online cyber security комп'ютерного, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкодію процесів моніторингу вхідних потоків malware-даних і упорядкування видаленням деструктивних компонентів.

4) *Нові* методи виявлення кіберзагроз, які здатні навчатися на великих даних, з метою виявлення кібератак, що можуть бути реалізовані у вигляді Інтернет посилань, поліморфних шкідливих програм (polymorphic malware) та троянських програм-шифрувальників, що займаються здирництвом: *метод атрибутно-орієнтованого впізнання Інтернет посилань* з використанням частотних шаблонів, що може провести оцінку атрибутів та відрізнити доброякісні, фішинг та шкідливі посилання; *метод детектування поліморфних шкідливих програм*, що дозволяє виявляти поліморфні шкідливі програми за допомогою аналізу хешів PE секцій та пошуку схожих секцій у бібліотеці шкідливого коду (після підтвердження детектування PE файлу нові хеші секцій додаються до бібліотеки); *метод пошуку криптопримітивів* у троянських програмах-шифрувальниках, що дозволяє виявити алгоритми шифрування, використовувани здирником, та дає можливість дешифрувати зашифровані файли користувача чи організації.

5) *Удосконалені* засоби захисту кіберпростору, які відрізняються використанням моделей і методів сигнатурно-логічного тестування атак, пошуку криптопримітивів у троянських програмах-шифрувальниках на основі використання алгоритмів машинного навчання, що дає можливість істотно скоротити час відновлення працездатності обчислювальної структури.

Практична реалізація результатів досліджень полягає у:

- тестуванні, верифікації і впровадженні розроблених програмних засобів перевірки, діагностування шкідливих програм і атак, що дає можливість виконувати їх моделювання із залученням існуючих додатків і malware бібліотек;
- програмній реалізації методу атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів, який відрізняється застосуванням апарату інтелектуального аналізу даних, що дає можливість визначати вірогідну оцінку небезпеки URL-адреси на основі атрибутів;
- програмній реалізації методу перевірки поліморфних шкідливих програм, який відрізняється інваріантністю до детермінізму сигнатур в код і урахуванням тільки контрольних сум Portable Executable (PE) секцій у виконуваних файлах, що дає можливість поліпшити продуктивність процедур діагностування деструктивних компонентів.

Результати дослідження у складі моделей, методів та інфраструктури впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акти про впровадження від 20.05.2019, 21.05.2019); у науково-виробничу діяльність компанії Design & Test Lab (довідка від 18.05.2019), у навчальний процес Blekinge Institute of Technology (BTH), Karlskrona, Sweden (лист ‘Statment of Reseach Results Impact on University Education Program’ від 29.05.2019).

СПИСОК ОПУБЛІКОВАНИХ РОБІТ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Список публікацій здобувача, в яких опубліковані основні наукові результати дисертації:

1. Hahanov V., Gharibi W., Litvinova E., Adamov A. Cyber Physical Computing for IoT-driven Services. – New York. USA. – Springer, 2018. 279p. [Chapter 2. Multiprocessor Architecture for Big Data Computing [Text] / V. Hahanov, W. Gharibi, E. Litvinova, A. Adamov. P. 21-41] (Springer, Scopus).
2. Carlsson A., Sokolianska I., Adamov A. Educating the Next Generation MSc in Cyber Security. – Sweden. – BTH, 2018. – 205 p. [Chapter: Education Challenges and Development of Advanced Malware Analysis Course [Text] / A. Adamov. P. 130-134].
3. Carlsson A., Sokolianska I., Adamov A. Educating the Next Generation MSc in Cyber Security. – Sweden. – BTH, 2018. – 205 p. [Chapter: The Cloud Threat Landscape [Text] / A. Carlsson, A. Adamov. P. 182-186].
4. Adamov A. Security risks and modern cyber security technologies for corporate networks [Text] / A. Adamov, V. Hahanov, W. Gharibi // Radioelektroniks and informatics. – 2010. – № 4. – P. 31–35. (Журнал реферується або індексується міжнародними базами Index Copernicus, Google Scholar, OECSF, OAJI, Scholar

Steer, SIS, Cyberleninka, CiteFactor, TIU Hannover, I2OR, Національної бібліотекою України ім. В. І. Вернадського).

5. Хаханов В.И. Инфраструктура анализа и информационной безопасности киберпространства [Текст] / В.И. Хаханов, С.В. Чумаченко, Е.И. Литвинова, А.С. Мищенко, А.С. Адамов // Радиоэлектроника и информатика. – 2011. – №2 (53). – С. 40-60. (Журнал реферується або індексується міжнародними базами Index Copernicus, Google Scholar, OECSP, OAJI, Scholar Steer, SIS, Cyberleninka, CiteFactor, TIU Hannover, I2OR, Національною бібліотекою України ім. В. І. Вернадського).

6. Адамов О.С. Блокчейн инфраструктура для захисту кіберсистем [Текст] / О.С. Адамов, В.І. Хаханов, С.В. Чумаченко, В.Г. Абдуллаєв // Радиоэлектроника и информатика. – 2018. – №4 (83). – С. 64-85. (Журнал реферується або індексується міжнародними базами Index Copernicus, Google Scholar, OECSP, OAJI, Scholar Steer, SIS, Cyberleninka, CiteFactor, TIU Hannover, I2OR, Національною бібліотекою України ім. В. І. Вернадського).

7. Адамов О.С., Хаханов В.І. Сигнатурно-кубітні методи розпізнавання деструктивних кодів [Текст] / О.С. Адамов, В.І. Хаханов // Радиоэлектроника и информатика. – 2019. – №1 (84). – С. 35-53. (Журнал реферується або індексується міжнародними базами Index Copernicus, Google Scholar, OECSP, OAJI, Scholar Steer, SIS, Cyberleninka, CiteFactor, TIU Hannover, I2OR, Національною бібліотекою України ім. В. І. Вернадського).

8. Adamov A. Analysis and Detection of Polymorphic Spyware [Text] / A. Adamov, A. Saprykin // Hakin9 Magazine. – 2013. – Vol. 8, № 01. – Issue 01/2013 (61). Warsaw: Software Press, 2013. – P. 6-11.

9. Adamov A.S., Hahanov V.I. A Method for the Attribute-based Detection of URLs Using Frequency Patterns [Text] / Вестник Государственного Инженерного Университета Армении. Серия: Информационные Технологии, Электроника, Радиотехника. – 2014. – Вып. 17, No2. – P. 59-66.

10. Сапрыкин А.С. Методика оценки убытков предприятия от вредоносных программ / Сапрыкин А.С., Бочарникова М.В., Адамов А.С. // Вестник национального технического университета "ХПИ" (Новое решение в современных технологиях). – 2009. – №8. – С. 58-64.

Результати, які засвідчують апробацію матеріалів дисертації:

11. Adamov A. Electronic System Level Models for Functional Verification of System-on-Chip / A. Adamov, K. Mostovaya, Syzonenko, I et al. // Proc. of International Conference “The Experience of Designing and Application of CAD Systems in Microelectronics” (CADSM). – Lviv-Polyana, Ukraine. – February 20-24, 2007. – P. 348–350. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

12. *Adamov A.* Transactional Data Analysis of Electronic System Level Models [Text] / *A. Adamov, V. Hahanov, D. Melnyk et al.* // Proc. of East-West Design and Test Symposium. September 7–10, 2007. Yerevan, Armenia. 2007. – P. 745–748.

13. *Adamov A.* Model of Source Code Analyzer for Hardware Description Languages [Text] / *Dmytro Melnyk, Sergei Zaychenko, Aleksandr Adamov, Vladimir Hahanov* // Proc. of East-West Design and Test Symposium (EWDTS). September 7–10, 2007, Yerevan, Armenia. – Yerevan, 2007. – P. 470–474.

14. *Hahanova I.V.* Transaction level model of embedded processor for vector-logical analysis [Text] / *Hahanova I.V., Obrizan V., Adamov A. et al* // Proc. of East-West Design and Test Symposium. September 2013. Rostov-on-Don, Russia. – P. 1–4. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

15. *Adamov A.* Data Mining Techniques for Functional Verification of SoC [Text] / *A. Adamov, R. Hwang, A. Gavrushenko* // Proc. of International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 2008. – Lviv-Polyana, Ukraine. – February 20-24, 2008. – P. 557-559.

16. *Adamov A.* The Problem of Trojan Inclusions in Software and Hardware [Text] / *A. Adamov, A. Saprykin* // Proc. IEEE East-West Design and Test Symposium (EWDTS'2009). – Moscow, Russia. – 18-21 September 2009. – P. 449-451. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

17. *Adamov A.* The problem of Hardware Trojans detection in System-on-Chip [Text] / *A. Adamov, A. Saprykin, D. Melnik* // Proc. CAD Systems in Microelectronics (CADSM 2009), – Lviv-Polyana, Ukraine. – 24-28 Feb 2009. – P. 178-179. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

18. *Adamov A.* Security risks in hardware: Implementation and detection problem [Text] / *Adamov A., Hahanov V.* // Proc. IEEE East-West Design and Test Symposium (EWDTS'2010). – Saint Petersburg, Russia. – September 17–20, 2010. – P. 425–427. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

19. *Adamov A.* A security model of individual cyberspace [Text] / *A. Adamov, V. Hahanov* // Proc. IEEE East-West Design and Test Symposium. – Sevastopol, Ukraine. – September 19–20, 2011. – P. 169–172. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

20. *Adamov A.* Discovering New Indicators for Botnet Traffic Detection [Text] / *A. Adamov, V. Hahanov, A. Carlsson* // Proc. of IEEE East-West Design & Test Symposium (EWDTS'2014). – Kiev, Ukraine. – September 26–29, 2014. – P. 281–285. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

21. *Hahanov V.* Structures for information retrieval in big data [Text] / *V. Hahanov, S. Chumachenko, E. Litvinova, A. Adamov et al* // Proc. of 13th International Conference Experience of Designing and Application of CAD Systems in Microelectronics (CADSM'2015). – Lviv, Ukraine. – 2015. – P. 70-75. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).

22. *Adamov A. A Sandboxing Method to Protect Cloud Cyberspace [Text] / A. Adamov, A. Carlsson // Proc. of IEEE East-West Design & Test Symposium (EWDTS'2015). – Batumi, Georgia. – September 27-30, 2015. – P. 180–183. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).*
23. *Adamov A. Android Ransomware: Turning CryptoLocker into CryptoUnlocker [Text] / A. Adamov // Proc. of the 25th Virus Bulletin International Conference. – Prague, Czech Republic. – 30 Sep – 2 Oct 2015. – P. 220-223.*
24. *Adamov A. Detecting targeted attacks in the cloud [Text] / A. Adamov // Proc. of the OpenStack Summit. – Vancouver, Canada. – 18-22 May 2015. 1 p.*
25. *Adamov A. Using Open Source Security Architecture to Defend against Targeted Attacks / Alexander Adamov, Dan Lambright // Proc. of the OpenStack Summit. – Austin, TX, US. – April 25-29, 2016. – 1 p.*
26. *Adamov A. Cloud incident response model [Text] / Alexander Adamov, Anders Carlsson // Proc. of IEEE East-West Design & Test Symposium (EWDTS'2016). – Yerevan, Armenia. – October 14–17, 2016. – P. 250–253. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).*
27. *Adamov A. The State of Ransomware. Trends and Mitigation Techniques [Text] / A. Adamov, A. Carlsson // Proc. of IEEE East-West Design & Test Symposium (EWDTS'2017). – Novy Sad, Serbia. – Sep 29 – October 2, 2017. – P. 121–128. (Входить до міжнародних наукометричних баз Scopus, IEEE Xplore).*
28. *Adamov A. Battlefield Ukraine: finding patterns behind summer cyber attacks [Text] / A. Adamov, A. Carlsson // Proc. of the 27th Virus Bulletin International Conference. Appendix: Last-minute presentations. – Madrid, Spain. – 4-6 Oct 2017. — P. 4-5.*
29. *Adamov A. Artificial Intelligence to Assist with Ransomware Cryptanalysis [Text] // Proc. of the 28th Virus Bulletin International Conference. – Montreal, Canada. – 3-5 Oct 2018. – P. 289-292.*
30. *Adamov A. Creating Ransomware Decryptors [Text] / A. Adamov // Proc. of 14th Cyber Security Conference UISGCON14. – Kyiv, Ukraine. – 2018. – P. 3.*
31. *Адамов А.С. Модель поиска вредоносного кода в программных продуктах [Текст] / А.С. Адамов, Д.А. Щербин, С.В. Чумаченко // Материалы 16-го Международного молодежного форума "Радиоэлектроника и молодежь в XXI веке". – Харьков, Украина. – 17-19 апреля 2012. – P. 121–123.*

АНОТАЦІЯ

Адамов О.С. Моделі і методи захисту кіберпростору на основі аналізу великих даних з використанням машинного навчання. – Дисертація на здобуття наукового ступеня кандидата технічних наук (доктора філософії) за спеціальністю 05.13.05 «Комп'ютерні системи та компоненти». – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2019.

Мета дослідження – істотне скорочення часу виявлення і блокування кібератак, спрямованих на кіберпростір суб'єкта, шляхом використання розроблених матричних моделей і логічних методів тестування, перевірки та діагностування за рахунок введення обчислювальної надмірності в інфраструктуру кіберпростору.

Наукова новизна результатів досліджень: 1) Удосконалено структурно-логічні моделі і методи перевірки кіберпростору для тестування і діагностування шкідливих компонентів, які відрізняються використанням методу дедуктивного паралельного аналізу обчислювальної системи для перевірки та діагностування malware. 2) Запропоновано нові методи синтезу еталонних логічних схем malware-функціональностей, які характеризуються використанням сигнатурно-кубітних структур, що дає можливість паралельно моделювати malware-driven великі дані для визначення належності поточного коду до існуючих деструктивних компонентів у malware бібліотеці. 3) Розроблено нову модель активного online cyber security комп'ютингу, яка характеризується сигнатурно-кубітним поданням інформації, що дає можливість підвищувати швидкодію процесів моніторингу вхідних потоків malware-даних і управління видаленням деструктивних компонентів. 4) Запропоновано новий метод атрибутно-орієнтованого розпізнавання URL-адрес з використанням частотних паттернів і метод перевірки поліморфних шкідливих програм на основі врахування контрольних сум Portable Executable секцій у виконуваних файлах і застосування апарату інтелектуального аналізу даних. 5) Удосконалено засоби захисту кіберпростору, які відрізняються використанням моделей і методів сигнатурно-логічного тестування атак, пошуку криптопримітивів у троянських програмах-шифрувальниках на основі використання алгоритмів машинного навчання, що дає можливість істотно зменшити час відновлення працездатності обчислювальної структури.

Результати дисертації у складі моделей, методів та інфраструктури впроваджені у навчальний процес Харківського національного університету радіоелектроніки (акти про впровадження від 20.05.2019, 21.05.2019); у науково-виробничу діяльність компанії Design & Test Lab (довідка від 18.05.2019), у навчальний процес Blekinge Institute of Technology (BTH), Karlskrona, Sweden (лист 'Statment of Reseach Results Impact on University Education Program' від 29.05.2019).

Результати дисертаційної роботи відображені в 31 друкованій праці, серед яких 13 публікацій у наукометричній базі Scopus.

Ключові слова: кіберпростір, великі дані, машинне навчання, логічний процесор, malware, кубітна модель, синтез тестів, вразливість, Cyber Security Computing.

АННОТАЦІЯ

Адамов А.С. Модели и методы защиты киберпространства на основе анализа больших данных с использованием машинного обучения. – Диссертация на соискание ученой степени кандидата технических наук (доктора философии) по специальности 05.13.05 «Компьютерные системы и компоненты». – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2019.

Цель исследования – существенное уменьшение времени обнаружения и блокирования кибератак, направленных на киберпространство субъекта, путем использования разработанных матричных моделей и логических методов тестирования, проверки и диагностирования за счет введения вычислительной избыточности в инфраструктуру киберпространства.

Задачи исследования:

1) Усовершенствовать структурно-логические модели и методы проверки киберпространства для тестирования и диагностирования вредоносных компонентов на основе использования дедуктивного анализа вычислительных систем.

2) Разработать сигнатурно-кубитные методы синтеза эталонных логических схем malware-функциональностей и параллельного моделирования malware-driven больших данных для определения принадлежности текущего кода к существующим деструктивным компонентам в malware библиотеке.

3) Разработать сигнатурно-кубитную модель активного online cyber security компьютеринга для мониторинга входных потоков malware-данных и управления процессом удаления деструктивных компонентов.

4) Усовершенствовать средства защиты киберпространства путем логического тестирования и диагностирования атак и вредоносных компонентов на основе использования алгоритмов машинного обучения.

5) Разработать метод атрибутно-ориентированного распознавания URL-адресов с использованием частотных паттернов и метод проверки полиморфных вредоносных программ на основе учета контрольных сумм Portable Executable секций в исполняемом файле и применения аппарата интеллектуального анализа данных.

6) Выполнить тестирование и верификацию разработанных программных средств тестирования, проверки и диагностирования вредоносных программ путем эмуляции атак на основе существующих malware библиотек.

Объект исследования – облачные и edge-computing технологии, основанные на высокопроизводительных дата-центрах, компьютерных гаджетах, системах и сетях, выполняющих сервисные функции хранения и анализа больших данных.

Предмет исследования – структурно-логические модели, методы, средства тестирования деструктивных компонентов, защиты индивидуального и коллективного сервис-компьютинга от киберугроз.

Научно-практическая задача – введение в инфраструктуру компьютерного пространства программной избыточности в форме моделей, методов и программных приложений для существенного уменьшения времени обнаружения и блокирования кибератак, направленных на киберпространство субъекта, путем использования разработанных матричных моделей и логических методов тестирования, проверки и диагностирования.

Сущность исследования – разработка моделей, методов и программных приложений для существенного уменьшения времени обнаружения и блокирования кибератак, направленных на киберпространство субъекта, путем использования разработанных матричных моделей и логических методов тестирования, проверки и диагностирования за счет введения вычислительной избыточности в инфраструктуру киберпространства.

Научная новизна результатов исследований:

1) Усовершенствованы структурно-логические модели и методы проверки киберпространства для тестирования и диагностирования вредоносных компонентов, которые отличаются использованием метода дедуктивного параллельного анализа вычислительной системы для проверки и диагностирования malware.

2) Предложены новые методы синтеза эталонных логических схем malware-функциональностей, которые характеризуются использованием сигнатурно-кубитных структур, что дает возможность параллельно моделировать malware-driven большие данные для определения принадлежности текущего кода к существующим деструктивным компонентам в malware библиотеке.

3) Разработана новая модель активного online cyber security компьютерного пространства, которая характеризуется сигнатурно-кубитным представлением информации, что дает возможность повышать быстродействие процессов мониторинга входных потоков malware-данных и управления удалением деструктивных компонентов.

4) Предложен новый метод атрибутивно-ориентированного распознавания URL-адресов с использованием частотных паттернов и метод проверки полиморфных вредоносных программ на основе учета контрольных сумм Portable Executable секций в исполняемые файлы с применением аппарата интеллектуального анализа данных.

5) Усовершенствованы средства защиты киберпространства, которые отличаются использованием моделей и методов сигнатурно-логического тестирования атак, поиска криптопримитивов в троянских программах-шифровальщиках на основе использования алгоритмов машинного обучения, что дает возможность существенно уменьшить время восстановления работоспособности вычислительной структуры.

Практическое значение полученных результатов исследований определяется:

б) Тестированием, верификацией и внедрением разработанных программных средств проверки, диагностирования вредоносных программ и атак, что дает возможность выполнять их моделирование с привлечением существующих приложений и malware библиотек.

7) Программной реализацией метода атрибутно-ориентированного распознавания URL-адресов с использованием частотных паттернов, который отличается применением аппарата интеллектуального анализа данных, что дает возможность определять вероятностную оценку опасности URL-адреса на основе его атрибутов.

8) Программной реализацией метода проверки полиморфных вредоносных программ, который отличается инвариантностью к детерминизму сигнатур в коде и учетом только контрольных сумм Portable Executable (PE) секций в исполняемом файле, что дает возможность улучшить производительность процедур диагностирования деструктивных компонентов.

Результаты диссертации в составе моделей, методов и инфраструктуры внедрены в учебный процесс Харьковского национального университета радиоэлектроники (акты о внедрении от 20.05.2019, 21.05.2019); в научно-производственную деятельность компании Design & Test Lab (справка от 18.05.2019), учебный процесс Blekinge Institute of Technology (BTH), Karlskrona, Sweden ('Statement of Research Results Impact on University Education Program' от 29.05.2019).

Результаты диссертационной работы отражены в 31 печатной работе: 3 раздела в зарубежных монографиях (из них 1 входит в наукометрическую базу Scopus), 7 статей (из них 5 – в научных журналах, включенных в «Перечень научных специализированных изданий Украины»; 2 статьи в международных научных журналах за рубежом; 4 статьи входят в международные наукометрические базы), а также в 21 международной научной конференции (из них 13 за рубежом, 12 входят в наукометрическую базу Scopus). Диссертант имеет 13 публикаций в наукометрической базе Scopus и индекс Хирша $h=3$.

Ключевые слова: киберпространство, большие данные, машинное обучение, логический процессор, malware, кубитная модель, синтез тестов, уязвимость, Cyber Security Computing.

ABSTRACT

Adamov Oleksandr Semenovich. Cyberspace protection models and methods based on big data analysis using machine learning. – Qualification scientific work as a manuscript. Thesis for the degree of candidate of technical sciences (Ph.D.) in specialty 05.13.05 "Computer systems and components". – Kharkov National University of Radio Electronics, Ministry of Education and Science of Ukraine, Kharkov, 2019.

The purpose of the study is to significantly reduce the time of detection and blocking of cyber attacks aimed at the cyberspace of the subject, using the developed matrix models and logical methods of testing, testing and diagnosing by introducing computational redundancy into the cyberspace infrastructure.

Objectives of the study:

1) Improve the structural and logical models and methods of testing cyberspace for testing and diagnosing malicious components based on the use of deductive analysis of computing systems.

2) Develop signature-qubit methods for the synthesis of standard logic circuits of malware-functionalities and parallel modeling of malware-driven big data to determine the belonging of the current code to the existing destructive components in the malware library.

3) Develop a signature-qubit model of active online cybersecurity computing for monitoring the input streams of malware-data and managing the process of removing destructive components.

4) Improve cyberspace security by logical testing and diagnosing attacks and malicious components based on the use of machine learning algorithms.

5) Develop a method of attribute-based URL recognition using frequency patterns and a method for testing polymorphic malware based on the accounting of the Portable Executable checksums of sections in the executable file and using the data mining apparatus.

6) Perform testing and verification of the developed software for testing, checking and diagnosing malware by emulating attacks based on existing malware libraries.

The object of the study is cloud and edge-computing technologies based on high-performance data centers, computer gadgets, systems and networks that perform the service functions of storing and analyzing big data.

The subject of the research is structural-logical models, methods, tools for testing destructive components, protection of individual and collective service computing from cyber threats.

The scientific and practical task is to introduce software redundancy into the infrastructure of the computing space in the form of models, methods and software applications to significantly reduce the time to detect and block cyber attacks aimed

at the subject's cyberspace by using the developed matrix models and logical methods of testing, testing and diagnosing.

The essence of the research is the development of models, methods and software applications to significantly reduce the time of detection and blocking cyberattacks aimed at the cyberspace of the subject, using the developed matrix models and logical methods of testing, checking and diagnosing by introducing computational redundancy into the cyberspace infrastructure.

The scientific novelty of research results:

1) Structural and logical models and methods for testing cyberspace for testing and diagnosing malicious components have been improved, which differ in using the method of deductive parallel analysis of a computing system to check and diagnose malware.

2) New methods for synthesizing reference logic circuits of malware-functionalities are proposed, which are characterized by the use of signature-qubit structures, which makes it possible to simulate malware-driven big data in parallel to determine whether the current code belongs to existing destructive components in the malware library.

3) A new model of active online cybersecurity computing has been developed, which is characterized by a signature-qubit representation of information, which makes it possible to increase the speed of monitoring the input malware-data streams and controlling the removal of destructive components.

4) Developing a method of attribute-based URL recognition using frequency patterns and a method for testing polymorphic malware based on the accounting of the Portable Executable checksums of sections in the executable file and using the data mining apparatus.

5) Cyberspace protection tools have been improved, which differ in the use of models and methods of signature-logic testing of attacks, the search for crypto-primitives in ransomware based on the use of machine learning algorithms, which makes it possible to significantly reduce the recovery time of the computing structure.

The practical significance of the results of research is determined by:

6) Testing, verification, and implementation of the developed software tools for checking and diagnosing malware and cyberattacks, which makes it possible to carry out their simulation with the assistance of existing applications and malware libraries.

7) Software implementation of the attribute-oriented URL recognition method using frequency patterns, which is distinguished by the use of the data mining apparatus, which makes it possible to determine the probabilistic risk assessment of the URL address based on its attributes.

8) Software implementation of the method for testing polymorphic malware, which is distinguished by the invariance to the determinism of signatures in the code and taking into account only the checksums of the Portable Executable (PE) sections

in the executable file, which makes it possible to improve the performance of diagnosing destructive components.

The reliability and validity of scientific results is confirmed by the accuracy of detecting and classifying examples of zero-day threats, including new versions of ransomware and advanced threats (Advanced Persistent Threats - APT); the positive feedback from academics and experts at international cybersecurity conferences such as the Virus Bulletin 2015 in Prague, Czech Republic, 2018 in Montreal, Canada, the OpenStack Summit 2015 in Vancouver, Canada and the OpenStack Summit 2015 in Austin, Texas, USA, IEEE East-West Design & Test Symposium (EWDTS'2017), 2017 in Novy Sad, Serbia.

The results of the thesis as part of the models, methods, and infrastructure are introduced into the educational process of the Kharkov National University of Radio Electronics (implementation certificates dated from 20.05.2019, 21.05.2019); in the research and production activities of the Design & Test Lab company (reference 18.05.2019), educational process of the Blekinge Institute of Technology (BTH), Karlskrona, Sweden ('Statment of Reseach Results Impact on University Education Program' from 29.05.2019).

The results of the dissertation work are reflected in 31 publications: 3 sections in foreign monographs (1 of them are in the Scopus science database), 7 articles (5 of them in scientific journals included in the "List of specialized scientific publications of Ukraine"; 2 articles in international scientific journals abroad; 4 articles are included in the international scientometric databases), as well as 21 international scientific conferences (13 of them abroad, 12 are included in the Scopus scientometric database). The author has 13 publications in the Scopus scientometric database and the Hirsch index $h = 3$.

Keywords: cyberspace, big data, machine learning, logical processor, malware, qubit model, test synthesis, vulnerability, Cyber Security Computing.